DOCUMENT RESUME

ED 139 607 SE 022 234

AUTHOR Crocker, Robert K.; And Others
TITLE An Experimental Study of Teacher Control in Sixth Grade Science Classes.
PUB DATE Mar 77
NOTE 57p.; Paper presented at the annual meeting of the National Association for Research in Science Teaching (50th, Cincinnati, Ohio, March 22-24, 1977); Contains occasional light type

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Educational Research; Elementary Education; *Elementary School Science; *Instruction; *Interaction Process Analysis; *Process Education; Research Methodology; Science Activities; Science Education; *Student Teacher Relationship; *Teacher Role
IDENTIFIERS Research Reports

ABSTRACT

The purpose of this study was to investigate the effects of two experimentally induced variations in teacher control in elementary science classes with respect to pupil achievement in science process skills, and pupil preferences and perceptions relative to the treatment variations. Interactions of treatment with pupil characteristics were also explored. Ten sixth-grade classes were exposed to an 18 week program consisting of activities from a process-based science curriculum. The program was taught by two teaching modes differing on a number of teacher behaviors related to the concept of control. Transcripts of videotapes of each class were coded using a modified Bellack system. Measured pupil characteristics included ability, personality, self-concept of ability, locus of control, socioeconomic status, previous science experience, and attitudes toward school and science. Criterion measures consisted of paper-and-pencil tests of science process achievement, an interview instrument, pupil preference of treatment, and a pupil choice task. Preliminary results indicated no overall treatment differences existed with respect to achievement but strong treatment effects occurred for pupil preference. (MH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED139607

AN EXPERIMENTAL STUDY OF TEACHER CONTROL IN SIXTH GRADE SCIENCE CLASSES

Robert K. Crocker, Roda P. Amaria, Douglas B. Sheppard,
Helen Banfield, Gary Rumboldt, and Glenn W. Clark
Institute for Educational Research and Development
Memorial University of Newfoundland
St. John's, Newfoundland, Canada

Paper Set Presented at the Annual Meeting of the
National Association for Research in Science Teaching
Cincinnati, Ohio
March, 1977

The research reported herein was supported by grants from the Humanities and Social Sciences Division of the Canada Council, and by the Institute for Educational Research and Development, Memorial University of Newfoundland. The findings and conclusions are those of the authors and not of the funding agencies. These findings are in preliminary form and should not be cited without permission of the authors.

SE 022 234

AN EXPERIMENTAL STUDY OF TEACHER CONTROL IN SIXTH GRADE SCIENCE CLASSES

I. BACKGROUND AND EXPERIMENTAL DESIGN

Robert K. Crocker, Roda P. Amaria, and Glenn W. Clark
Memorial University of Newfoundland

The study reported in this and succeeding papers in this set was designed to address the general question of teacher control in the specific context of elementary science classes. The classes under investigation were characterized by direct pupil investigation of science problems, including pupil data collection and interpretation. Important surface features of the classroom setting included pupils working in small groups each with its own set of apparatus, relatively large amounts of pupil-pupil interaction, and teacher interaction with individuals and groups as well as with the class as a whole. The setting was thus similar to that which might be associated with most of the major elementary science curriculums.

The concept of control, as applied to this setting, essentially involves what might be referred to as the degree to which the teacher "orchestrates" the activities of the pupils. The degree of moment-to-moment teacher direction of activity, the degree to which the teacher gave specific instructions for apparatus manipulation and data collection, the degree to which the pupils were required to pursue a common goal with common procedures, the amount of teacher-class versus teacher-group interaction, and so on are elements of control

in the sense used here. For purposes of investigation, the concept was given more specific operational meaning through a number of teacher behaviours derived from a modified version of the Bellack classroom observation system (Bellack, et al., 1966). Two experimental treatments, labelled low control (LC) and high control (HC) were defined in terms of these categories.

The general objective of the study was to investigate the influence of treatment of pupil classroom behaviours, achievement, preference for treatment, and perceptions of the goals and procedures of the lessons being taught, and to explore possible interactions between pupil characteristics and treatment. A secondary objective was to examine possible differences between teachers in their implementation of the treatments and to determine whether treatment-teacher interactions exert an influence on outcome variables. The objectives may thus be expressed in terms of a treatment main effects hypothesis and two interaction hypotheses treatment by pupil characteristics and treatment by teachers, with each hypothesis being tested with respect to three types of dependent variables, classroom behaviours, pupil achievement, and pupil perceptions and preferences. The results reported in the present paper set pertain only to the latter two classes of variables. Classroom behaviours are currently being coded and will be reported in due course.

Rationale of the Study

In a general sense, the study might be considered as an attempt to investigate teaching at the level of "strategy". That is, the study attempted to define variations in teacher behaviour at a level at which the teacher might be expected to have some control of the variations. This level may be distinguished from the "tactical" level at which behaviours are in terms of categories such as "praise" or "questioning" and from what might be termed the "methods" level, in which behaviour variations are more at the level of curriculum, materials, and the like, which are typically outside the teacher's control. The level at which the study aims is therefore a level at which, given a particular curriculum, a particular school setting, and a particular group of pupils, the teacher might be expected to make a choice of the manner in which the lesson is to be conducted.

The rationale for the focus on teacher control derives from Barker's behaviour setting construct (Barker, 1963) and from Bernstein's concept of strong and weak framing (Bernstein, 1971). More specifically, Barker has developed the notion that an individual's perception of the goals of the setting may influence his behaviour in that setting. Thus high and low teacher control may lead to different pupil perceptions of goals. For example, it might be hypothesized that in a high control setting the pupil may be more likely to perceive the goal in terms of meeting immediate teacher requirements, thus reducing the opportunity to deal with the cognitive demands of the problem at hand. Grannis (1973),

following Bernstein, has proposed that classroom behaviours may be influenced by whether a particular setting is congruent or incongruent in terms of teacher versus pupil control of a number of parameters of the setting. It might be argued that high teacher control of lesson pacing, apparatus arrangements, data collection, type of results to be sought, and so on is incongruent with pupil use of apparatus and high pupil-pupil interaction. The nature of the behaviour differences and the influence on outcomes of congruent and incongruent settings is a subject for empirical investigation.

The behaviour setting construct admits the possibility that an individual's behaviour is a function of that individual's characteristics as well as of parameters of the setting. A more formal expression of this concept is given by Hunt (1971; Hunt and Sullivan, 1974) in the context of the question of matching teaching styles to pupil characteristics. Hunt's analysis is based on an elaboration of Lewin's paradigm $B=f(P,E)$ (behaviour is a function of the person and the environment). The same basic concept, of course, provides the basis for the increased emphasis on aptitude-treatment interactions in studies of teaching.

At a more practical level, it might be argued that the variations in teacher behaviour defined by the treatments in this study are within the limits that would normally be open to the teacher involved in curriculums such as S-APA, SCIS, or ESS. In such curriculums, in fact, the teacher has little control of content or of materials. The teacher is, however, somewhat free to determine lesson pacing, the specific

form of interactions with pupils, and other variations within the lesson itself. A curriculum such as ESS perhaps has one further degree of freedom in that different groups of pupils could conceivably work on different aspects of the same topic. This aspect of teacher choice is not included in the definitions of treatments in this study.

Related Research

Although a large body of research exists on the general question of control, the problem has been viewed from many different perspectives, thus making comparisons difficult. For example, the large number of studies generated by the Flanders Interaction Analysis Category System (FIAC) have perceived control in terms of teacher "indirectness". Kounin (1970) has investigated control in terms of discipline and group management. Behaviour modification studies have been concerned with control of deviant behaviour. Control of classroom organization, grouping, materials use, movement and the like have been treated in some of the large scale classroom observation studies (e.g. Soar and Soar, 1973; Brophy and Evertson, 1974; Stallings and Kaskowtiz, 1974; MacDonald and Elias, 1976).

Results from studies in the Flanders tradition seem to suggest that increased "indirectness" is associated with increased pupil achievement and attitude scores. Kounin's research indicates that a number of "group management" variables are useful as predictors of pupil work involvement and deviant behaviour. Soar and Soar (1973) found nonlinear

relationships between variables associated with control and pupil achievement, suggesting that the optimum amount of teacher control may be somewhere between the extremes of the continum. Stallings and Kaskowitz, (1974) suggest that greater "openness" is associated with higher scores on affective types of measures but that greater control is associated with achievement on cognitive measures. Brophy and Evertson (1974) report interactions between control variables and socioeconomic status, with pupils in low SES schools benefitting generally from greater control.

A number of more recent experimental studies have a more direct bearing on the problem of interest here than do the studies cited above. Shymansky and Matthews (1974) report an experiment on two learning conditions in fifth grade science. The learning conditions, labelled "teacher structured" and "student structured" were defined in terms of a number of observable teacher behaviours related to teacher directiveness. Results showed that student classroom behaviours differed markedly under the two treatments. Surprisingly, almost no behaviour differences occurred as a function of student characteristics. Interactions between student ability and treatment were also generally nonsignificant. With science process achievement as the criterion, the student structured treatment led to significantly higher scores. In terms of the present study, it is interesting to note that the amount of time spent by the teacher in interacting with individual pupils correlated negatively with the amount of time spent by the student in activities of his own design. This suggests the need for a closer examination of teacher-student interaction.

A similar, but smaller scale study by Crocker, Bartlett, and Elliott (1976) yielded results in the opposite direction from those of Shymansky and Matthews. Students in a more highly structured learning environment performed better on a science process achievement test. In this study, a repeated measures design was employed so that it was possible to obtain a measure of preference for treatment. Again the results favoured the structured treatment. Interactions with student neuroticism (anxiety) were also found for both the achievement and preference criteria.

Differences in the results of these two studies may be due to differences in methodology. Since the Shymansky and Matthews study was of much longer duration, it is possible that improved performance in an unstructured environment occurs only after a period of orientation to this environment. Also, it is not clear that the definitions of the treatments were sufficiently parallel in the two studies to permit direct comparison.

An experiment that even more closely parallels the present one has been conducted at the Stanford Centre for Research and Development in Teaching (Program on Teaching Effectiveness, 1976). In this experiment, a three-way factoral design, based on two levels each of teacher Structuring, Soliciting, and Reacting (Bellack et al, 1966), was employed. Teacher behaviours on each of the factors were controlled by using scripted lessons. In general, the results indicate that low soliciting was more effective in inducing achievement on test items for which the teacher was the only source of information. High structuring and high reacting were also

more effective than their low counterparts. Thus, the treatment combination of low soliciting-high structuring-high reacting was the most effective in this criterion. No differences between treatments occurred, however, on an essay achievement test and on an attitude scale. A generalized regression analysis indicated, as might be expected, that student ability accounted for much more of the variance in achievement than did treatment effects. It was also found that ATI effects accounted for about the same proportion of total variance as treatment effects. No conclusions about specific sources of interactions were reported. An interesting feature of the study was the use of path analysis to determine the degree to which student perceptions of treatments may have mediated the treatment effects. Such mediating effects were found to occur for the structuring and reacting treatment but not for the soliciting treatment.

The present study incorporates a number of design improvements over the earlier Crocker, Bartlett, and Elliott study. Also, the duration of each treatment should have been sufficiently long to remove this as a possible explanation for the contradiction in results of that and the Shymansky and Matthews study. In terms of the Program on Teaching Effectiveness experiment, to the extent that our high control treatment and the PTE high structuring-low soliciting-high reacting treatment reflect the underlying notion of what Rosenshine (1976) refers to as "direct instruction", there is reason to predict that, at least for achievement, the high control treatment should be more effective. The use of the repeated measures design again allows a measure of preference to be

developed and permits what amounts to two replications of the experiment in a single study.

Experimental Design

The original sample consisted of eleven grade six classes in six schools within the Avalon Consolidated School District in St. John's, Newfoundland. One of these classes was subsequently dropped from the main analysis when it was found that pupils had already completed some of the material included in the experimental program. This class was then used for a number of other purposes in the study, particularly in reliability studies on some of the instruments.

Four experienced elementary teachers attached to the project staff were given a one-week training program on the treatments and the general outline of lesson content. One of these teachers was employed full-time to act as coordinator for lesson preparation, distribution of materials, testing, and related tasks. The remaining teachers were part-time and were free to be assigned to classes in accordance with the experimental requirements. Regular classroom teachers were replaced by these project teachers for all science periods during the eighteen-week course of the experiment. In this case, the loss of generalizability stemming from the use of special teachers was offset by the substantial gain in experimental control of being able to assign teachers to classes in a relatively balanced fashion.

Because of the relatively small number of classes, it was decided to employ a repeated measures design in which

all classes were exposed to both treatments. At the beginning of the experimental period, five classes were assigned to low control and six to high control. After nine weeks, the treatments were reversed and the teachers reassigned to new classes. The assignment of teachers was random, with the restriction that each teacher have at least one class in each treatment. Scheduling difficulties imposed some constraints on teacher assignments. However, these constraints were not considered related to any of the experimental variables and were therefore not considered departures from randomness.

These design considerations are summarized in Table I. The occurrence of six classes in one treatment and four in the other resulted from the loss of the eleventh class as mentioned above. On the whole, the design permitted the analysis of treatment main effects and treatment by student characteristics interactions, free from confounding due to teacher and class effects. Some confounding does remain in assessing treatment-teacher interactions within a round. However, this confounding can be removed to some degree by treating the two rounds as replications, with teachers being assigned to different classes in the two rounds.

Data Sources

Table II gives a summary of variables used in the study, along with instruments used and, where appropriate, reliability coefficients. Individual classroom behaviour variables are not listed because of the large number of these variables and because only teacher behaviours are of interest in this report.

TABLE I

EXPERIMENTAL DESIGN

ROUND	TREATMENT	
	HC	LC
1	1, 2 A	3 A
	4 B	5, 9 C
	10 C	7 D
	6, 8 D	
2	3 A	1, 4 A
	9 B	8 B
	7 C	6 C
	5 D	2,10 D

KEY: A - D TEACHERS

1 -10 CLASSES

TABLE II

SUMMARY OF VARIABLES AND DATA SOURCES

VARIABLE	SOURCE OF DATA	RELIABILITY COEFFICIENT	RELIABILITY METHOD
PUPIL BACKGROUND			
AGE	SCHOOL RECORDS	-	
SEX	SCHOOL RECORDS	-	
SOCIOECONOMIC STATUS	BLISHEN SCALE	-	
YEARS OF SCIENCE EXPERIENCE	PUPIL QUESTIONNAIRE	-	
TYPE OF SCIENCE EXPERIENCE	PUPIL QUESTIONNAIRE	-	
PUPIL ABILITIES			
NON-VERBAL INTELLIGENCE	RAVEN PROGRESSIVE MATRICES	.83 - .93	TEST-RETEST
BASIC SKILLS	CANADIAN TEST OF BASIC SKILLS	.87 - .97	SPLIT-HALF
- LANGUAGE SKILLS	CANADIAN TEST OF BASIC SKILLS		
- MATH SKILLS	CANADIAN TEST OF BASIC SKILLS		
- WORK STUDY SKILLS	CANADIAN TEST OF BASIC SKILLS		
- COMPOSITE	CANADIAN TEST OF BASIC SKILLS		

TABLE II CONTINUED

SUMMARY OF VARIABLES AND DATA SOURCES

VARIABLE	SOURCE OF DATA	RELIABILITY COEFFICIENT	RELIABILITY METHOD
READING ABILITY			
- VOCABULARY	GATES-MACGINTIE	.89	SPLIT-HALF
- COMPREHENSION	READING TEST	.95	SPLIT-HALF
PERSONALITY TRAITS			
EXTRAVERSION	JUNIOR EYSENCK PERSONALITY INVENTORY	.70 - .80	SPLIT-HALF AND
NEUROTICISM	JUNIOR EYSENCK PERSONALITY INVENTORY		TEST-RETEST
DOMINANCE - SUBMISSIVENESS	CHILDREN'S PERSONALITY QUESTIONNAIRE	.74 - .32	SPLIT-HALF AND PARALLEL FORMS
DEPENDANCY	DEPENDANCE PRONENESS SCALE (FLANDERS ET AL, 1960)	.68 (ORIGINAL AUTHOR) .91 (ELLIOTT 1973)	TEST-RETEST
SELF CONCEPT OF ABILITY	MODIFIED BROOK-OVER SCALE	.72	TEST-RETEST
LOCUS OF CONTROL	INTELLECTUAL ACHIEVEMENT RESPONSIBILITY SCALE		
ATTITUDES TOWARDS SCHOOL	SEMANTIC DIFFERENTIAL	.83	TEST-RETEST

TABLE II CONTINUED

SUMMARY OF VARIABLES AND DATA SOURCES

VARIABLE	SOURCE OF DATA	RELIABILITY COEFFICIENT	RELIABILITY METHOD
ATTITUDE TOWARDS SCIENCE	SEMANTIC DIFFERENTIAL	.74	TEST-RETEST
SCIENCE PROCESS ACHIEVEMENT	20 ITEM MULTIPLE CHOICE TESTS OF SCIENCE PROCESS OF MEASURING, CONTROLLING VARIABLES, AND INTERPRETING DATA	.69 (PRETEST) .74, .82 (POSTTEST)	TEST-RETEST
PUPIL PREFERENCES AND PERCEPTIONS			
- PREFERENCE SCALE	22 ITEM AGREE-DISAGREE SCALE BASED ON VARIABLES USED TO DEFINE TREATMENTS	.71	TEST-RETEST
- CHOICE TASK	PUPILS ASKED TO CHOOSE TEACHING MODE FOR A SCIENCE ACTIVITY AND TO CARRY OUT THE ACTIVITY UNDER THIS MODE	-	

TABLE II CONTINUED

SUMMARY OF VARIABLES AND DATA SOURCES

VARIABLE	SOURCE OF DATA	RELIABILITY COEFFICIENT	RELIABILITY METHOD
- INTERVIEW SCHEDULE	25 ITEM STRUCTURED INTERVIEW ON GOALS OF ACTIVITIES, TEACHER ROLE, PUPIL RELATION WITH PARTNER, ETC.	-	
CLASSROOM BEHAVIOURS (INDIVIDUAL CATEGORIES NOT LISTED TO CONSERVE SPACE).	VIDEOTAPES OF 30 CLASS SESSIONS. MODIFIED BELLACK OBSERVATION SYSTEM, EMPHASIS ON TEACHER CONTROL CATEGORIES.	.50 - .99 (DEPENDING ON DIMENSION AND METHOD)	INTER-CODER AGREEMENT, VARIOUS CODING TECHNIQUES

In order to avoid subjecting pupils to a heavy testing load at the beginning of the study, data on pupil characteristics were gathered at different times throughout the study, with the restriction that a particular instrument be administered to all classes at about the same time. Data on basic skills (CTBS) were obtained from the school district which was engaged in a testing program while the study was in progress. These data were thus acquired at no cost to the study and with minimal disruption of experimental class sessions. The Gates-MacGintie reading test had been administered in two of the schools prior to the beginning of the study. Rather than administer a repeat of this test in these schools, the test norms were used to project the obtained scores to the time at which the remaining tests were administered.

Implementation of Treatments

Table III gives the variables on which teachers were expected to differ in implementing the high and low control treatments. In general, the differences are in the desired direction. Exceptions occur for teacher responding and reacting, which show essentially no differences, and pupil reporting individual results, which shows a difference in the opposite direction from that predicted. In the latter case, however, the large inter-lesson variance within a treatment renders the result unreliable. In any case, this form of pupil behaviour was perhaps not under teacher control to a sufficient degree to show predictable effects.

TABLE III

COMPARISON OF TEACHER BEHAVIOURS IN HIGH AND LOW CONTROL TREATMENTS

VARIABLE	HIGH CONTROL	LOW CONTROL
TEACHER-CLASS INTERACTION	876*	419
TEACHER-GROUP INTERACTION	671	1,095
PROPORTION OF PUPIL TALK DURING TEACHER-CLASS INTERACTION	.06	.10
PROPORTION OF PUPIL TALK DURING TEACHER-GROUP INTERACTION	.24	.29
TEACHER STRUCTURING	453	369
TEACHER SOLICITING	539	650
TEACHER RESPONDING	89	84
TEACHER REACTING	229	227
RATIO OF COMMANDS TO TOTAL SOLICITS	.31	.21
RATIO OF REQUESTS TO TOTAL SOLICITS	.69	.79
PUPILS REPORTING INDIVIDUAL RESULTS	76	46
TEACHER SUBSTANTIVE-LOGICAL DISCOURSE	187	230
PUPIL SUBSTANTIVE-LOGICAL DISCOURSE	34	55
AVERAGE WAIT TIME (SEC)	1.5	1.9

*UNITS ARE TOTAL SENTENCE UTTERED IN A SAMPLE OF APPROXIMATELY 130 MINUTES LESSON TIME FOR EACH TREATMENT. SAMPLE INCLUDES 13 LESSONS WITH ALL TEACHERS INCLUDED AT LEAST ONCE IN EACH TREATMENT. RATIOS AND WAIT TIME COMPUTED FOR SOME SAMPLE DATA.

The observed difference for teacher soliciting, while in the required direction, was not as strong as expected. This suggests that the strength of the soliciting move in defining teacher role, as has been well established in other studies, is such that changes cannot easily be brought about, even through training. The specific character of the solicits as commands or requests, however, did differ in the direction in which teachers were expected to behave.

AN EXPERIMENTAL STUDY OF TEACHER CONTROL IN SIXTH GRADE SCIENCE CLASSES

II. SCIENCE PROCESS ACHIEVEMENT

Helen Banfield, Robert K. Crocker, and Roda P. Amaria
Memorial University of Newfoundland

Objective

The general rationale for the study has been described in the preceeding paper. The specific objective of this part of the study was to develop measures of science process achievement and to compare experimental treatments and explore interactions between treatment and pupil characteristics with respect to pupil performance on these measures.

Instruments

Three twenty-item multiple choice tests of science processes were developed for use as pretest and first and second round posttests respectively. Items for these tests were adapted from the competency measures of Science-A Process Approach, the AAAS Science Process Instrument, and the Test of Science Processes by Tannenbaum (1971). Items were restricted to the processes of measuring, controlling variables, and interpreting data since these were the processes emphasized in the experimental lessons.

An initial pool of forty items for each test was used to prepare four parallel forms. Items were designed to reflect both the immediate content of the lessons and more general applications of the processes of interest. Item analysis procedures were used to obtain data on item difficulty and discrimination. The tests were also used in a pilot study to check reading level.

Test-retest reliability was determined using one sixth grade class in a school in which the same material used in the study was being taught but which was not used in the main analysis (class eleven of the original sample). Reliability of the pretest was established at .69 while the posttests yielded reliabilities of .74 and .82 respectively.

Method

The experimental design was as described in the first paper. The pretest was administered during the first week of the study. Posttests were given at the end of each nine-week treatment round. The four parallel forms of each test were given to random samples of pupils in each class to reduce bias due to repeated responding to the same items and to reduce the effect on mean scores of specific items. The latter precaution was taken because of the relatively small number of items on each test.

For purposes of analysis, the two treatment rounds were considered as replications and analyzed separately. Raw posttest scores were first adjusted for pupil ability, using the composite CTBS score as the covariate. CTBS rather than

pretest was used because it was found to correlate more highly with posttests. Pretest means were quite low, probably reflecting lack of previous exposure of subjects to the teaching of processes. It was therefore considered possible that the degree of randomness in responses to the pretest was greater than that suggested by the reliability coefficient, which had been determined using a class which was known to have had at least some previous exposure to processes.

Residual posttest scores were subjected to a series of three-way analyses of variance, with treatment, teacher, and each pupil characteristic, in turn, as factors. Pupil characteristics were dichotomized or tricotomized as appropriate. The requirement that interactions be analyzed led to the use of individual pupil scores rather than class means as the unit of analysis. This procedure was justified on the grounds that, in the present context, treatments could be considered to act on the individual more than on the class as a unit, since many interactions took place at the individual level rather than the class level. Also, the normal coherence of the class unit was disrupted to a degree by the use of special teachers.

Because of the large number of main and interaction effects to be tested (some 140 statistical tests in all), it was inevitible that using the normal .05 level of significance would yield a number of false positives. Because of this, and because two replications were available, a modified decision rule was adopted whereby the .10 level of significance was accepted for a single test but where a result was considered

significant only if the probability was less than that level for both replications. The consequence of this decision rule was to reduce the probability of Type I errors to the order of .01, without the attendant increase in the probability of Type II errors that would exist in a single replication.

Results

Table I gives the probabilities associated with the F ratio for the various tests of interest. As can be seen, only the treatment by sex and the treatment by dependency interactions were significant in terms of the decision rule adopted. Tables II and III give the cell means corresponding to these interaction effects. It is apparent from these tables that, in spite of stability in the existence of overall interactions, these interactions do not show consistent patterns in their detail. For example, while the first round results show that females do particularly poorly in low control, no similar pattern emerges in the second round.

TABLE I

PROBABILITIES OF F FOR MAIN EFFECTS AND INTERACTIONS, POSTTEST RESIDUALS AS CRITERION

SOURCE OF VARIANCE*	PROBABILITY	
	ROUND 1	ROUND 2
TREATMENT	NS*	NS
TEACHER	NS	NS
TREATMENT X TEACHER	NS	NS
TREATMENT X SEX	.002	.05
TREATMENT X SES	.99	.02
TREATMENT X EXTRAVERSION	.05	.23
TREATMENT X NEUROTICISM	.37	.07
TREATMENT X DOMINANCE	.01	.99
TREATMENT X DEPENDENCY	.001	.04
TREATMENT X LOCUS OF CONTROL	.06	.99
TREATMENT X SELF-CONCEPT	.09	.99
TREATMENT X ATTITUDE TO SCHOOL	.18	.31
TREATMENT X ATTITUDE TO SCIENCE	.99	.29

*ALL PROBABILITIES BASED ON 3-WAY (TREATMENT X TEACHERS X PUPIL CHARACTERISTICS) ANOVA TO EACH PUPIL CHARACTERISTIC SEPARATELY. PROBABILITIES FOR TREATMENT AND TEACHER MAIN EFFECT VARIED WITH DIFFERENT STUDENT CHARACTERISTICS BUT WERE GENERALLY NONSIGNIFICANT.

TABLE II

CELL MEANS FOR TREATMENT X SEX INTERACTION

TREATMENT		SEX M	SEX F
ROUND 1	HC	.05	.05
	LC	.10	-.24
ROUND 2	HC	.14	-.14
	LC	.08	-.08

*MEANS ARE STANDARDIZED POSTTEST RESIDUALS.

TABLE III

CELL MEANS FOR TREATMENT X DEPENDENCY INTERACTION

TREATMENT		DEPENDENCY H	M	L
ROUND 1	HC	.04*	.24	-.15
	LC	.22	-.45	.07
ROUND 2	HC	-.01	-.01	.12
	LC	-.25	.05	.16

*MEANS ARE STANDARDIZED POSTTEST RESIDUALS.

AN EXPERIMENTAL STUDY OF TEACHER CONTROL IN SIXTH GRADE SCIENCE CLASSES

III. PUPIL PREFERENCE

Douglas B. Sheppard, Robert K. Crocker, and Roda P. Amaria
Memorial University of Newfoundland

Objective

The purpose of this phase of the study was to investigate treatment effects and trait-treatment interactions using a measure of pupil preference for treatment and a measure of pupil choice of treatment as dependent variables. The general hypothesis under investigation pertained to interaction effects in that it was expected that, independently of any possible treatment effects on achievement, certain pupil characteristics, particularly personality traits, would influence pupil liking for one or the other treatment. The basis for this expectation lies in the nature of teacher behaviours in the two treatments as expanding or restricting pupil freedom to carry out activities in a manner of their own choosing. In addition, the treatment behaviours were such as to increase or decrease the amount of pupil-pupil interaction and hence the opportunity for pupils to engage in interpersonal relationships.

Method

Two separate criterion measures were used for pupil preference. The first was a twenty-two item agree-disagree scale consisting of one positively and one negatively

worded item for each of eleven behaviours corresponding to those used in the original operational definitions of the treatments. For example, pupils were asked whether they agreed or disagreed with statements such as "I would prefer more (or fewer) teacher directions than are given in science class" or "I would prefer more (or less) teacher interaction with the whole class".

This instrument was administered at the end of each treatment so that in each case pupil responses were made from the perspective of the treatment just completed. In deriving a score on the instrument, it was necessary to scale responses according to the treatment, and to match responses on the positive and negative item of each pair. Four possible response patterns could occur for a given item pair (agree with both, agree with negative item and disagree with positive item, agree with positive item and disagree with negative item, or disagree with both). By eliminating the agree with both as an inconsistent response and by summing over all item pairs, it was possible to derive a score representing the individual's liking for the treatment just completed.

Test-retest reliability for this instrument was established at .71.

The second criterion measure was designed to assess the same underlying trait but was intended to be more task oriented. At the end of each treatment, pupils were asked to choose a lesson to be carried out in either of two teaching modes, corresponding approximately to the two treatments.

Essentially the choice was between an activity in which the purpose was defined and apparatus supplied but no further direction given and the same activity in which a worksheet was supplied giving detailed directions for the activity.

Data from this task consisted simply of recording the choice made by each pupil. Once the choice was made, pupils were required to carry out the activity in the mode chosen, in order to ensure that the choice was based on the expectation of actually having to perform in that mode.

Data from the preference scale were analyzed using a series of two-way analyses of variance with treatment and each student characteristic, in turn, as the factors. A treatment by teacher analysis was also carried out. Responses on the choice task were crosstabulated by treatment and by each pupil characteristic variable and the chi-square test applied to the resulting contingency tables. As in the analysis of achievement scores, the decision rule was to require a probability level of .10 or less for both replications in order for a result to be considered significant.

Results

Table I gives the probabilities of F for the relevant main and interaction effects. The highly significant treatment effect is complicated by the presence of a number of significant two-way interactions. It should be noted, however, that the treatment main effect is independent of the interactions in the sense that the variance due to interaction effects was removed before testing the main effect.

The presence of the interactions led to an examination of the tables of cell means to determine whether the results were consistent in detail across the two replications. For this purpose, it was decided that the comparison of trends in the two replications was more useful than carrying out multiple statistical comparisons of cell means within a round. That is, given the overall significance of an interaction, the result was considered significant if the tendency in the cell means was the same for both rounds rather than if a particular pair of cell means were significantly different.

The following is a brief summary of the findings of the cell means inspection:

1. The direction of the treatment main effect was in favour of the low control treatment for both rounds.

2. Because classes taught by different teachers were of different ability, the treatment by ability interaction is confounded with the treatment by teacher interaction within a round. However, the treatment by ability interaction was quite consistent for the two rounds, with high ability students exhibiting a stronger preference for high control and low ability pupils showing the opposite tendency, even though teachers had been interchanged. Similarly, the treatment-teacher interaction, which showed one teacher clearly obtaining higher preference scores than expected in low control, was consistent for the two rounds. It therefore seems reasonable to conclude that these are stable and independent effects.

3. The treatment-SES interaction showed consistent effects with high SES students showing greater preference for low control and low SES students greater preference for high

TABLE I

PROBABILITIES OF F FOR MAIN AND INTERACTION EFFECTS, PREFERENCE CRITERION

SOURCE OF VARIANCE	PROBABILITIES	
	ROUND 1	ROUND 2
TREATMENT	<.001	<.001
TEACHER	<.001	<.001
TREATMENT X TEACHER	<.001	<.001
TREATMENT X SEX	.99	.99
TREATMENT X CTBS COMPOSITE	<.001	<.001
TREATMENT X SES	.007	<.001
TREATMENT X EXTROVERSION	.008	.07
TREATMENT X NEUROTICISM	.99	.10
TREATMENT X DOMINANCE	.002	.003
TREATMENT X DEPENDENCY	.99	.99
TREATMENT X LOCUS OF CONTROL	.99	.35
TREATMENT X SELF-CONCEPT	.04	.23
TREATMENT X ATTITUDE TO SCHOOL	.30	.99
TREATMENT X ATTITUDE TO SCIENCE	.04	.99

control. This effect parallels the treatment-ability interaction since, surprisingly, a negative (small but significant) correlation was found between ability and SES.

4. The treatment-extraversion interaction was not stable over the two rounds. The tendency for low extraverts to more strongly prefer high control than high extraverts was consistent, whereas the pattern for low control showed a reversal from the first round to the second.

5. Finally, the treatment-dominance interaction showed a clear and strong pattern, with high dominants preferring low control and low dominants preferring high control.

The distributions of pupil choice of high and low control tasks by treatment are shown in Table II. It is apparent that pupil choice was the independent of treatment. This result was perhaps to be expected since the choice task was not directly tied to the treatment the pupils had just completed. However, the row totals suggest that pupils showed a strong tendency to select the low control task. (χ^2 values for the observed distributions, based on the expectation of equal numbers of choices for each type of task, are significant at the .01 and .001 levels for Rounds 1 and 2 respectively). It is possible that the relatively stronger tendency shown in Round 2 was due to pupils having had the benefit of exposure to both alternatives, but this cannot be directly confirmed from the data.

TABLE II

FREQUENCIES OF PUPIL CHOICE OF HIGH OR LOW CONTROL TASK BY TREATMENTS

CHOICE		TREATMENT H	TREATMENT L	TOTAL	χ^2
ROUND 1	CHOSE HIGH	85(83)	53(55)	138	.22 N.S.
	CHOSE LOW	110(112)	78(76)	188	
	TOTAL	195	131	326	
ROUND 2	CHOSE HIGH	31(34)	59(56)	90	.58 N.S.
	CHOSE LOW	93(90)	141(144)	234	
	TOTAL	124	200	324	

These results are consistent with the findings for the preference scale. The question remains, however, whether pupil choices were related to pupil characteristics. To answer this question, a series of contingency tables of pupil choice by pupil characteristic variable were developed. A similar table was developed for pupil choice by teacher. Non-independence of pupil choice and a particular pupil characteristic was interpreted as evidence of interaction between choice and that characteristic. This analysis could thus be regarded as paralleling that for the preference scale.

A summary of the findings for the χ^2 tests appear in Table III (contingency tables have been omitted in the interest of brevity). It can be seen that the results for teachers and for pupil dominance parallel those for the preference scale. In addition, a strong sex effect and a moderately strong dependency effect appear for the choice task. The ability, extraversion, and SES effects for the preference scale are not replicated for the choice task. The finding for self-concept of ability is perhaps worth noting, even though it is marginally nonsignificant according to the decision rule, because it does show a consistent pattern for both criterion measures.

TABLE III

SUMMARY OF CHI-SQUARE TESTS FOR PUPIL CHOICE OF HIGH OR LOW CONTROL TASK

VARIABLE	ROUND 1			ROUND 2		
	χ^2	DF	P	χ^2	DF	P
TEACHER	9.68	3	<.025	8.60	3	<.05
SEX	12.86	1	<.001	15.77	1	<.001
CTBS COMPOSITE	3.81	2	NS	4.16	2	NS
SES	3.50	2	NS	2.88	2	NS
EXTRAVERSION	4.07	2	NS	3.23	2	NS
NEUROTICISM	3.98	2	NS	2.88	2	NS
DOMINANCE	15.35	2	<.001	19.21	2	<.001
DEPENDENCY	6.17	2	<.05	7.12	2	<.05
LOCUS OF CONTROL	2.04	2	NS	5.43	2	<.10
SELF CONCEPT	4.44	2	NS*	9.67	2	<.01
ATTITUDE TO SCHOOL	2.22	2	NS	1.68	2	NS
ATTITUDE TO SCIENCE	2.14	2	NS	1.65	2	NS

*MARGINAL, P≃.11

AN EXPERIMENTAL STUDY OF TEACHER CONTROL IN SIXTH GRADE SCIENCE CLASSES

IV. PUPIL INTERVIEW DATA

Gary Rumboldt, Glenn W. Clark, Robert K. Crocker
and Roda P. Amaria
Memorial University of Newfoundland

Objectives

The first general hypothesis under investigation in this phase of the study was that pupils would tend to perceive the high control treatment in different terms from the low control treatment. More specifically, it was expected that pupils under high control would perceive the activities more in procedural terms and in terms of teacher requirements and directions, while those in low control would hold perceptions that were oriented towards the process nature of the activities and towards freedom to pursue the activities in ways other than that defined by the teacher. As in other aspects of the study, it was also expected that pupil responses would be related to pupil characteristics.

A second objective of this part of the study was to investigate the relationships between pairs of pupils assigned to work together during the activities. The hypothesis here related less to treatments than to the issue of the dynamics of the pupil-pair group structure and the matching of pupils in such groups.

Accordingly, the design of the study required the

development and application of a pupil interview schedule designed to yield data on pupil perceptions of the purposes of the activities, of the degree of teacher control employed, and of relationships between members of the pupil-pairs which constituted the basic grouping arrangement in the activities. Pupil interviews were chosen as an alternative to other techniques such as rating scales or checklists in order to allow for open-ended responses and to permit probing of the reasons for specific pupil responses.

Originally, it had been intended that the interviews also include a component whereby pupils would complete an assigned task designed to measure their performance in the specific process objectives of the lesson just completed. This would have provided an independent measure of achievement. However, because of time limitations on the interviews, this objective could not be achieved.

Method

The interview schedule was designed to provide a standard format whereby interviewers could ask questions about 1) whether the pupils perceived the activities in process, content, procedural, or other terms, 2) how the pupils perceived the frequency, clarity, and sufficiency of teacher directions for the conduct of the activities, 3) pupil perceptions of the amount of freedom available in carrying out the activities, and 4) the pupil's relationship with his partner. In general, the interviewer was required to categorize responses in terms of a set of predefined categories, while

allowing for the possibilities of responses that did not conform to these categories. For most questions, pupils were asked why they gave a particular response.

On four occasions throughout the experimental sequence, eight pupils from each class were selected as interview subjects. The interviews took place immediately following a science activity, and questions were framed in the context of that activity. During the first round, interviews were recorded on audio tape for discussion among interviewers and to permit reliability checks.

The preliminary results given here are based on cross-tabulations of the item responses by treatment and various pupil characteristics. The chi-square statistic has been used in tests of significance. As in previous analyses, consistency across interview occasions, as well as level of significance of an individual statistical test, was part of the decision rule.

Results

Individual interview questions tabulated over four rounds of interviews yielded a large number of contingency tables. For this reason, the results are presented here in summary form for rounds one and four rather than in the form of individual statistical tests. The summary in Table I is based on clusters of items having the common themes listed. Differences due to treatment, teacher, sex, SES and personality were found for items dealing with perceptions of the nature of the problem under investigation in the lessons. Items on

teacher structuring of activity and teacher versus pupil role yielded responses that would have been expected to be more closely associated with high control than with low control. The teacher was generally viewed as playing a strong leadership role in the activities for round one, independently of treatment. Over a number of weeks, after being exposed to Science lessons through activity approach, in round four there was a shift among low control towards pupil role being perceived as more important only on the grounds that the pupils did all the work. Yet in both rounds the teacher is perceived as the source of procedures and judge of results.

A summary of responses as related to pupil characteristics appears in Table II.

TABLE I

SUMMARY OF PUPIL RESPONSES TO INTERVIEW ITEM CLUSTERS BASED ON ROUNDS ONE AND FOUR

ITEM CLUSTER	RESULT
1. PERCEPTION OF LESSON IN CONTENT, PROCESS, PROCEDURAL, OR OTHER TERMS.	IN BOTH ROUNDS LOW CONTROL PERCEIVED LESSONS IN CONTENT TERMS. HIGH CONTROL IN ROUND ONE PERCEIVED MORE IN PROCEDURE TERMS, WHEREAS IN ROUND FOUR MORE IN CONTENT TERMS. FOR BOTH ROUNDS RELATIVELY FEW PUPILS PERCEIVED LESSONS IN PROCESS TERMS. FOR THE TWO ROUNDS THERE WERE DIFFERENCES DUE TO TEACHERS IN THE WAY THE PUPILS PERCEIVED LESSONS. ALSO DIFFERENCES WERE FOUND DUE TO PUPIL CHARACTERISTICS, (GIVEN SEPARATELY IN TABLE II).
2. TEACHER STRUCTURING	GENERALLY ALL PUPILS, IRRESPECTIVE OF TREATMENTS, SEX, SES, ABILITY OR PERSONALITY DIFFERENCES, PERCEIVED TEACHER AS THE SOURCE OF PROCEDURE AND JUDGE OF RESULTS. PUPILS ABOUT EVENLY DIVIDED ON WHETHER THEY COULD HAVE CARRIED

TABLE I CONTINUED

SUMMARY OF PUPIL RESPONSES TO INTERVIEW ITEM CLUSTERS BASED ON ROUNDS ONE AND FOUR

ITEM CLUSTER	RESULT
	OUT ACTIVITY WITH LESS TEACHER DIRECTION.
3. TEACHER VERSUS PUPIL ROLE	IN ROUND ONE TEACHER ROLE PERCEIVED AS MORE IMPORTANT THAN PUPIL ROLE BECAUSE OF NEED TO FOLLOW TEACHER DIRECTIONS AND IMPORTANCE OF OBTAINING RESULTS TEACHER WANTED. IN ROUND FOUR HIGH CONTROL PERCEIVED TEACHER ROLE MORE IMPORTANT, WHEREAS LOW CONTROL PUPILS PERCEIVED PUPIL ROLE AS MORE IMPORTANT BECAUSE THEY FELT THAT THEY DID ALL THE WORK (χ^2, P=.004). IF GIVEN THE CHOICE, MOST PUPILS WOULD HAVE DONE ACTIVITY THE SAME WAY AS THE TEACHER ASKED.

TABLE I CONTINUED

SUMMARY OF PUPIL RESPONSES TO INTERVIEW ITEM CLUSTERS BASED ON ROUNDS ONE AND FOUR.

ITEM CLUSTER	RESULT
4. TIME	NO TREATMENT, SEX, ABILITY OR PERSONALITY DIFFERENCES IN PERCEIVED ADEQUACY OF TIME TO CARRY OUT ACTIVITY. EXCEPT WITH ONE TEACHER, IN A SCHOOL WHERE THE CLASS PERIOD WAS OF SHORTER DURATION (30 MINS. INSTEAD OF 40 MINS.).
5. PUPIL-PARTNER RELATIONSHIPS	PUPIL PAIRS GENERALLY WORKED WELL TOGETHER. PUPILS TENDED TO GIVE PARTNER MORE CREDIT, OR THOUGHT THAT BOTH SHARED THE WORK BUT VERY RARELY GAVE CREDIT TO HIMSELF FOR PERFORMANCE IN THE ACTIVITY.
	NO TREATMENT, SEX, ABILITY OR PERSONALITY DIFFERENCES FOUND.

TABLE II

SUMMARY OF PUPIL RESPONSES TO INTERVIEW ITEM CLUSTERS BY PUPIL CHARACTERISTICS BASED ON ROUNDS ONE AND FOUR

PUPIL CHARACTERISTIC	RESULT
SEX	FOR BOTH ROUNDS, MALE AND FEMALE PUPILS DID NOT DIFFER IN THEIR PERCEPTION OF LESSON IN CONTENT TERMS. IN THEIR PERCEPTION OF LESSON IN TERMS OTHER THAN CONTENT, PROCESS OR PROCEDURE THERE WERE MORE FEMALES THAN MALES. THERE WAS NO DIFFERENCE AMONG MALES AND FEMALES ON OTHER CLUSTERS.
SOCIO-ECONOMIC STATUS	MORE PUPILS OF MEDIUM SES PERCEIVED LESSONS IN EITHER PROCEDURE, OR CONTENT TERMS. IN ROUND ONE, MORE LOW SES PUPILS PERCEIVED LESSONS IN TERMS OTHER THAN CONTENT, PROCESS OR PROCEDURE. NO DIFFERENCES DUE TO SES FOR OTHER CLUSTERS.

TABLE II CONTINUED

SUMMARY OF PUPIL RESPONSES TO INTERVIEW ITEM CLUSTERS BY PUPIL CHARACTERISTICS BASED ON ROUNDS ONE AND FOUR

PUPIL CHARACTERISTIC	RESULT
ABILITY	RELATIVELY MORE PUPILS PERCEIVED LESSONS IN CONTENT TERMS, INDEPENDENTLY OF ABILITY. OF THOSE WHO PERCEIVED LESSON IN PROCESS TERMS, GREATER NUMBERS WERE OF HIGH ABILITY RATHER THAN OF LOW ABILITY FOR BOTH ROUNDS. THERE WAS A TREND IN ROUND FOUR FOR MORE HIGH ABILITY PUPILS TO PERCEIVE PUPIL ROLE AS MORE IMPORTANT THAN TEACHER ROLE. NOT MUCH DIFFERENCE ON ABILITY FOR OTHER CLUSTER ITEMS.
EXTRAVERSION	GENERALLY MORE PUPILS PERCEIVED LESSONS IN CONTENT TERMS. OF THE ONES WHO PERCEIVED LESSONS IN PROCESS TERMS THEY WERE MORE EXTROVERTS THAN INTROVERTS, IN BOTH ROUNDS. FOR TEACHER VERSUS PUPIL ROLE THERE WAS A TREND IN ROUND FOUR

TABLE II CONTINUED

SUMMARY OF PUPIL RESPONSES TO INTERVIEW ITEM CLUSTERS BY PUPIL CHARACTERISTICS BASED ON ROUNDS ONE AND FOUR

PUPIL CHARACTERISTIC	RESULT
	FOR HIGH EXTROVERTS TO PERCEIVE PUPIL ROLE MORE IMPORTANT THAN THE TEACHER ROLE.
NEUROTICISM	GENERALLY MORE PUPILS PERCEIVED LESSONS IN CONTENT TERMS FOR BOTH ROUNDS. THOSE WHO PERCEIVED LESSON IN PROCEDURE TERMS THERE WERE MORE HIGH ANXIOUS THAN LOW ANXIOUS. THERE WAS A TREND IN ROUND FOUR FOR HIGH ANXIOUS TO PERCEIVE TEACHER ROLE AS MORE IMPORTANT, WHEREAS LOW ANXIOUS PERCEIVED PUPIL ROLE AS MORE IMPORTANT.
DOMINANCE-SUBMISSIVENESS	FROM AMONG THE PUPILS WHO PERCEIVED LESSONS IN PROCESS TERMS, THERE WERE MORE DOMINANT ONES. THERE WAS NO DIFFERENCE ON OTHER CLUSTERS.

TABLE II CONTINUED

SUMMARY OF PUPIL RESPONSES TO INTERVIEW ITEM CLUSTERS BY PUPIL CHARACTERISTICS BASED ON ROUNDS ONE AND FOUR

PUPIL CHARACTERISTIC	RESULT
DEPENDENCY, LOCUS OF CONTROL, SELF-CONCEPT	NO DIFFERENCES DUE TO THESE PUPIL CHARACTERISTICS FOR CLUSTER ITEMS.
YEARS AND TYPE OF SCIENCE EXPERIENCE	IN ROUND ONE, IT WAS UNUSUAL TO FIND THAT MORE PUPILS WITH VERY LITTLE PREVIOUS SCIENCE EXPERIENCE AS COMPARED TO PUPILS WITH MORE SCIENCE EXPERIENCE PERCEIVED LESSONS EITHER IN PROCESS TERMS, OR IN TERMS OTHER THAN CONTENT, PROCESS OR PROCEDURE. MORE PUPILS WITH PREVIOUS EXPERIENCE OF SCIENCE FROM TEXT BOOKS PERCEIVED LESSONS IN PROCEDURE TERMS.

AN EXPERIMENTAL STUDY OF TEACHER CONTROL IN SIXTH GRADE SCIENCE CLASSES

V. SUMMARY AND DISCUSSION

Robert K. Crocker and Roda P. Amaria
Memorial University of Newfoundland

SUMMARY

The general purpose of the study was to investigate the effects of two experimentally induced variations in teacher control in elementary science classes with respect to pupil achievement in science process skills, and pupil preferences and perceptions relative to the treatment variations. Interactions of treatment with pupil characteristics were also explored.

Ten grade six calsses were exposed to an eighteen week program consisting of activities from a process based science curriculum. The program was taught by two teaching modes differing on a number of teacher behaviours related to the concept of control. Each class was taught for nine weeks in each mode. Teaching was carried out by four teachers assigned to the project. This permitted teacher training, random assignment of teachers to classes, and balanced assignment of teachers to treatments (all teachers taught both treatments). The use of both treatments in all classes reduced confounding due to unique class-treatment effects by permitting an analysis in terms of two replications of the treatments.

Each class was videotaped on three occasions during the term, yielding approximately eight taped sessions for each teacher. Transcripts of the tapes were coded using a

modified version of the Bellack classroom observation system. Observed teacher behaviours were used to monitor the implementation of treatments and to investigate differences between teachers. Pupil behaviours are being coded and are to be used as dependent variables in a later phase of analysis.

Pupil characteristic variables included measures of ability (Raven Progressive Matrices, Canadian Test of Basic Skills, Gates-McGintie Reading Test), personality (Junior Eysench, Personality Inventory, Dependency Scale, Dominance Scale), self-concept of ability, locus of control (IAR), socioeconomic status (Blishen Scale), sociometric status, previous science experience, and attitudes towards school and science. Criterion measures consisted of paper-and-pencil tests of science process achievement, items from a pupil interview instrument (dealing with pupil perceptions of the activities), a rating scale for pupil perception of treatment from which was derived a measure of preference, a pupil choice task, and a number of pupil behaviour categories from the classroom observation instrument.

Preliminary results reported in the preceeding papers indicate that no overall treatment differences existed with respect to achievement but that strong treatment effects occurred for pupil preference. Interaction effects of treatment with sex and dependency occurred for achievement but these were unstable over the two replications. Somewhat stronger and more consistent interactions occurred for the two preference criteria. Pupil interviews yielded only one source of difference between treatments, that of pupil perceptions of the nature of the activities as process, content or procedural.

DISCUSSION

Achievement

The absence of stable patterns of cell means in the observed interaction effects, in spite of overall statistical significance, sheds doubt on the generalizability of the achievement results. One possible explanation for the change for the two rounds might be the fact that in Round 2 subjects had had the benefit of exposure to two alternatives. That is, the Round 2 results may have been due to a combination of Round 1 and Round 2 influences. This multiple treatment effect is present in any design of a repeated measures nature. This explanation is reinforced slightly by a secondary analysis in which a repeated measures ANOVA was used to compare performance over the two rounds by the same pupils. In this analysis, the nature of the interactions seemed influenced by the order in which the treatments were presented.

While it is tempting to rely on the repeated treatments effect to account for the observed inconsistencies, it must be noted that the ultimate aim of experiments of this sort is to find main effects and interactions that are sufficiently generalizable to form the basis for decision making on grouping or on the matching of teaching styles to student characteristics. If interaction effects are as context specific as the repeated treatments argument suggests, then such effects are of little value in practice.

Preference

Although it seems reasonable to state that strong

and consistent treatment effects, in favour of low control, did occur for the two preference criteria, this statement does oversimplify the situation somewhat. The treatment-teacher interaction shows that the treatment effect was much stronger for some teachers than for others. It is, of course, possible (even likely) that different teachers applied the treatment in different ways, in spite of the attempt to define specific treatment behaviours and to train teachers. If this is so, a further analysis of treatment variables by teacher should reveal the sources of difference. An even more likely prospect, however, is that teachers could have differed on behaviours relevant to the treatments but not included in the operational definitions. In particular, it can be hypothesized that a relevant variable might be the degree to which a treatment is congruent with the teacher's "natural" teaching style. There is some evidence, in fact, that some teachers were more at home in high control and others in low control. It may be possible, in subsequent analysis, to associate subjective judgements of congruence of teaching style and treatment with teacher behaviours and with pupil response on the criterion instruments. Ideally, this would require the coding of teacher behaviours other than those used in differentiating treatments. Videotape records exist which could permit such an analysis.

Again, it would appear that pupil preferences for high and low control are strongly related to pupil characteristics. The fact that, except for dominance, the specific interactions do not replicate over the two preference criteria again raises the question of the usefulness of such effects.

The situation here, however, is somewhat different from that for the achievement criterion because consistent effects over replications did occur. The interactions seem to be criterion specific rather than occasion specific. It is perhaps not surprising that pupils might respond in one way when asked for preference for a specific behaviour (such as more or fewer teacher directions) and when asked to carry out a task involving variations in the same type of behaviour. For example, a highly dependent pupil may perceive few teacher directions as desirable but, when faced with a task in which few directions are given, may feel unable to cope with the perceived desirable situation. In this sense, the choice task, in spite of its unsophistication as a measurement technique, may be a more valid measure than the preference scale with its complex scoring scheme. Certainly, for example, the grouping of pupils in a class on the basis of their choice of teaching mode, and the prediction of such choices from pupil characteristics, is an entirely feasible procedure.

A broader issue is, however, raised by the discrepancy between interactions for the two criteria of preference. That is the question of whether pupils should be matched to teaching styles on the basis of interactions or whether an attempt should be made to change the pupil characteristics themselves. Again using the example of the dependent pupil, it might be argued that in the long term, it is more desirable to make the pupil more independent than to assign him to a teaching mode which reinforces his dependency. The analysis of discrepancies between desired teaching modes and actual teaching modes chosen might form an interesting starting point for research on this issue.

Interview Data

The interview data seems to suggest that pupil perceptions of the treatments were generally closer to what might have been expected for high control than for low control. That is, pupils appear to have responded to interview items as if both treatments involved a relatively high degree of teacher control. Although evidence from the observation of teacher behaviours and from the preference scale indicate that the treatments were, indeed, different, it is possible that neither treatment involved such an extreme of low control that pupils would have perceived teacher directions as inadequate or pupil role as more important than teacher role. This problem of implementation of low control was expected to occur from the outset. Nevertheless, the investigators were reluctant to place inordinate pressure on teachers to apply an extreme version of low control since this would have clearly been inconsistent with any reasonable teaching style that these teachers would have been expected to use under ordinary classroom conditions.

Further Analysis

It is stressed that the findings presented here represent only the first stage of analysis of the data. The coding of teacher and pupil classroom behaviour remains incomplete. In particular, the issue of whether teachers differ in their specific ways of implementing the treatments remains to be explored. There also seems to be reason to argue that treatment differences may be more pronounced for classroom behaviour than for achievement.

Further analysis on the achievement and preference criteria is underway. This analysis uses a generalized regression technique to assess the relative effects of prior ability, treatments, and interactions and to remove the limitation in assessing interactions of blocking on the pupil characteristics variables.

Finally, it seems desirable to examine the possible mediating effects of classroom behaviour and pupil perception on pupil achievement. It is proposed that such effects be analyzed when the coding of pupil behaviours is complete.

REFERENCES

Barker, R.G., 1968. Ecological Psychology. Stanford, Calif., Stanford University Press.

Bellack, A.A., Hyman, R.T., Smith, F.L., Jr., and Klichard, H.M., 1966. The Language of the Classroom. New York, Teachers College Press.

Bernstein, B., 1971. Class, codes, and control. London, Routledge and Kegan Paul.

Brophy, J.E. and Evertson, C.M., 1974. The Texas Teacher Effectiveness Study: Presentation of Nonlinear Relationships and Summary Discussion. Austin, Texas, Research and Development Centre for Teacher Education, University of Texas.

Crocker, R.K., Bartlett, K.R., and Elliott, H.G., 1976. A comparison of structured and unstructured modes of teaching science process activities. Journal of Research in Science Teaching, 13, 267-274.

Grannis, J.G., 1973. Columbia Classroom Environment Project. (Extracts from final report). USOE Contract No. OEC-71-0593. New York, Teachers College, Columbia University.

Hunt, D.E., 1971. Matching Models in Education. Toronto, Ontario Institute for Studies in Education, Monograph Series No. 10.

Hunt, D.E., and Sullivan, E.V., 1974. Between Psychology and Education. Hinsdale, Illinois, Dryden Press.

McDonald, F.J., 1976. Beginning Teacher Evaluation Study, Phase II: Summary Report. Princeton, New Jersey: Educational Testing Service.

Program on Teaching Effectiveness, 1976. A Factorially Designed Experiment on Teacher Structuring, Soliciting, and Reacting. (R & D Memorandum No. 147). Stanford, California, Stanford Centre for Research and Development in Teaching.

Rosenshine, B., 1976. Classroom Instruction. In N.L. Gage (ED.), The Psychology of Teaching Methods. 77th yearbook of the National Society for the Study of Education. Chicago, University of Chicago Press.

Shymansky, J.G., and Matthews, C.C., 1974. A comparative laboratory study of the effects of two teaching patterns on the behaviour of students in fifth grade science. Journal of Research in Science Teaching, 11, 157-168.

Soar, R.S., and Soar, R.M., 1973. Classroom Behaviour, Pupil Characteristics, and Pupil Growth for the School Year and the Summer. Gainesville, Florida, Institute for Development of Human Resources, University of Florida.

Stallings, J.A., and Kaskowitz, D.H., 1974. Follow Through Classroom Observation Evaluation, 1972-73. Menlo Park, California, Stanford Research Institute.